An Architecture for Almost All-Optical Nodes Based on Wavelength Converters

Konstantinos Yiannopoulos, Kyriakos Vlachos and Emmanuel (Manos) Varvarigos

***Abstract*—We present an almost all-optical node architecture suitable for on-the-fly packet and burst switching without losses. The operation of the node is based on wavelength converters for mapping the incoming to the outgoing links, and for intra-node contention resolution. The node can be built out of commercially available equipment, and is easily scalable, with respect to the number of its incoming and outgoing links, by simple addition of components.**

***Index Terms*—Optical networks, optical packet switching, almost all-optical nodes, wavelength converters, Mach-Zehnder interferometers.**

I. Introduction

OPTICAL networks have seen a tremendous increase in line rates during the last decade, particularly due to the progress in WDM and OTMD technologies, which enabled the transmission of an aggregate of more than 10 Tbps [1] and the single channel transmission of 1.28 Tbps [2] over a single fiber. Despite the aforementioned progress in transmission speeds, fully exploiting the potential of optical networks requires optical nodes of similar bandwidth to be deployed. The design of high bandwidth packet-switched optical nodes is restricted by opto-electronic and electro-optic conversion, which is common practice in currently installed networks. Although format conversion is particularly desirable, since the processing capabilities of electronics exceed those of photonics, its bandwidth is limited to a few tens of GHz and is not expected to allow for channel rates over 40 Gbps [3], [4]. In order to overcome this difficulty, almost all-optical node architectures have been proposed [5], [6]. In almost all-optical nodes only the control plane information is converted into electronic format, and are therefore processed in the node control plane, whereas the rest of the data remain in the optical domain and traverses the optical node. Format conversion for the control plane data would still be feasible even in the case of ultra-high line rates, for instance by setting the control data rate to be lower than the line rate or by stuffing them periodically into the stream. The almost all-optical approach combines the best of both worlds, photonics and electronics, since it minimizes format conversion, allows for reliable electronics-based control planes, and maintains ultra-high bandwidth for the core of the node.

Apart from achieving ultra-high bandwidth in optical nodes, it is of equal importance to utilize the available bandwidth in the most effective fashion, allocating resources on a per need basis rather than statically. This in turn means that circuit switching, which currently dominates WDM networks, should be altered to optical packet or burst switching [7], [8]. Optical packet and burst switching are expected to fully utilize the bandwidth of future optical networks, however there are still open issues preventing them from full deployment. The main shortcoming is the unavailability of reliable random access optical memories, thus, optical packets cannot be forwarded on a store-and-forward basis, as in electronic IP nodes, and an on-the-fly forwarding approach must be considered. However, on-the-fly forwarding has a strong probability of causing packet losses due to contention inside the core of the optical node, for instance when two optical packets require the same node output simultaneously. As a result, the optical node architecture should allow for contention resolution, thus lossless operation.

In this paper we describe an almost all-optical node architecture suitable for lossless on-the-fly optical packet and burst forwarding. The node optical core comprises of an all-optical switching matrix, which performs spatial mapping between the node input and output ports on a packet-by-packet (or burst-by-burst) basis, cascaded by a scheduler, which comprises of optical fiber lines of programmable delay, to allow for contention resolution. Both the switching matrix and the scheduler are based on wavelength converters and wavelength selective filters, so as to fully exploit the parallelism of spatial and wavelength separation in the node. Space and wavelength parallelism offers the lossless (contention-free) operation of the node under certain burstiness properties of the transferred packets, while in the same time enables the minimization of the number of the components that are required to build the node, therefore its cost. The node is simple to build out of commercially available equipment, such as MZI based wavelength converters, arrayed waveguide gratings and standard single

Manuscript submitted February 10, 2006. This work is supported by the European Union and the Greek Government under EPEAEK Programme Pythagoras II.

Authors are with the Department of Computer Engineering and Informatics, University of Patras, Rio 26500, Greece (phone: + 30 2610 996990; fax: +30 2610 969007; e-mail: giannopu@ceid.upatras.gr). They are also with the Research Academic Computer Technology Institute (RACTI), Rio 26500, Greece

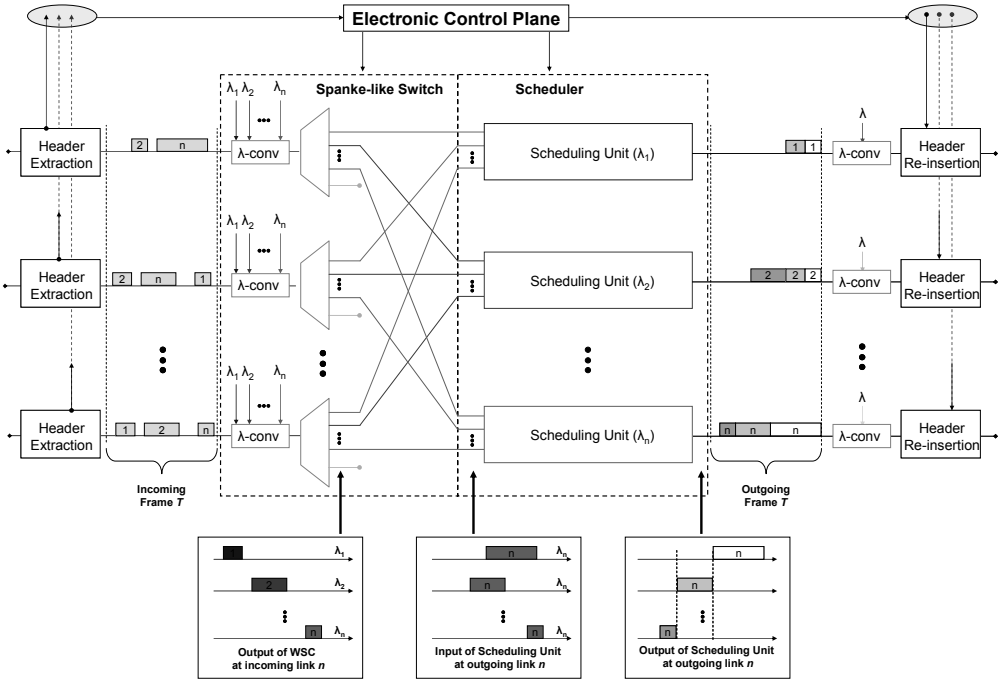

Fig. 1. Architecture of the almost all-optical node. Incoming packets are directed to the outgoing links by means of wavelength converters (λ-conv) and wavelength selective couplers. Conention resolution is enforced in the scheduling units.

mode optical fiber, and may be easily extended, as far as the number of input and output ports are concerned, by simply adding more components.

The rest of the paper is organized as follows: in Section II we discuss in detail the architecture of the node, and we emphasize on the switch and scheduler implementation. We also discuss a scheduling algorithm that provides lossless operation even for bursty traffic, provided that certain burstiness properties are met. In Section III we focus on the operation of the wavelength converter when deployed for packet and burst switching, and comment on cascadability issues that arise from the node architecture. Finally, Section IV concludes the paper.

II. Node Architecture

The architecture of the node having n incoming links and n outgoing links is shown in Fig. 1. Upon arrival, the headers of the optical packets in all incoming links are read and directed to the electronic plane of the node [9]. The electronic plane processes the header information and generates the appropriate control signals that will set the state of the node for the packet payload to transverse the node core on-the-fly. The node core comprises a Spanke-like space switch and a programmable time-delay scheduler. The Spanke-like switch has n input ports, equal to the number of the incoming links. In each input port a wavelength converter (λ-conv in Fig. 1) operates on a per packet basis, performing wavelength conversion on the payloads of the packets. The decision of the wavelength to convert to is taken by the control plane, depending on the outgoing link that each packet requires access to. The packet payloads are directed to their respective outgoing links through wavelength selective couplers (WSCs), which make physical connections between the input and output ports of the switch. The output ports of the switch are grouped in n clusters, each comprising n branches. The total number of branches is nxn, contrary to the number of outgoing links which is only n. Redundancy is introduced in the switch output ports so that even if all incoming packets request the same output port simultaneously, their payloads are directed to different branches, and consequently no blocking (packet loss) occurs in the switch. Possible contention at the node outgoing links, when two or more packet payloads arrive at the same time and at the same output cluster of the switch, are resolved inside the scheduling units, one per outgoing link, that comprise the scheduler. The scheduling units enforce contention resolution by delaying the packets at their inputs relative to one another. We will discuss in more detail the scheduling units' structure and an algorithm for avoiding contention in Subsection II.A. After exiting the scheduler subsystem, the packet payloads are assigned with a common wavelength, which corresponds to their outgoing link, and they have new headers inserted.

A. Contention Resolution Algorithm

Contention resolution is feasible inside the scheduling unit, provided that the packets that simultaneously arrive at it are delayed relative to each other, and therefore never coincide at the node outgoing link [10]. To this end, we assume that the time axis is divided in time slots and a single data unit (DU) occupies the respective time slot. There is a maximum number of slots the scheduling unit may resolve contention within, which we will denote as frame size T. This approach is valid

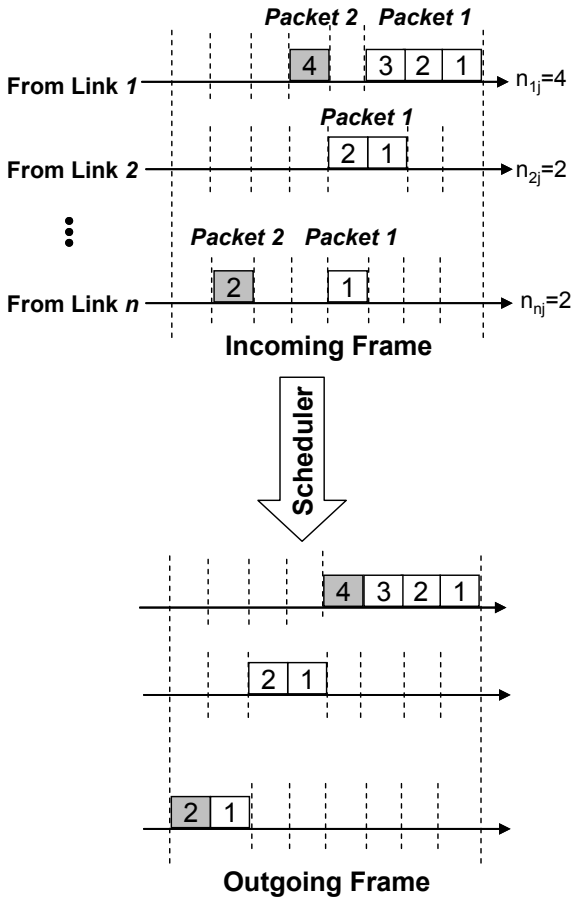

Fig. 2. Collision resolution in outgoing link *j*, based on relative time delay of packet payloads. Each payload equals a multiple of Data Units. The numbers inside the packets correspond to the order of the respective Data Unit.

for both optical packet switching and optical burst switching, where the packet or burst consists of a varying number of DUs while its maximum length is equal to the frame size. We further assume that from each incoming link *i* the total traffic (in DUs) requesting the outgoing link *j* is equal to n_{ij}. Provided that the total traffic in link *j* does not exceed the link capacity

$$\sum_{i=1}^{n} n_{ij} \le T, \qquad j = 1,.., n \tag{1}$$

or equivalently that the traffic is *(n,T)*-smooth, we may focus on a single DU *A* that requests access to the outgoing link *j*. Assuming that *A* is the r_A^{th} DU that arrives in the scheduler from incoming link *i* and heads for outgoing link *j*, we define the order of *A* to be equal to $r_A \in \{1,\ldots,n_i\}$. Contention is avoided by requesting that *A* is placed in slot

$$y_A = \sum_{l=1}^{i-1} n_{lj} + r_A - 1, \;\; y_A \in \{0,..., T-1\} \tag{2}$$

of the outgoing link frame.

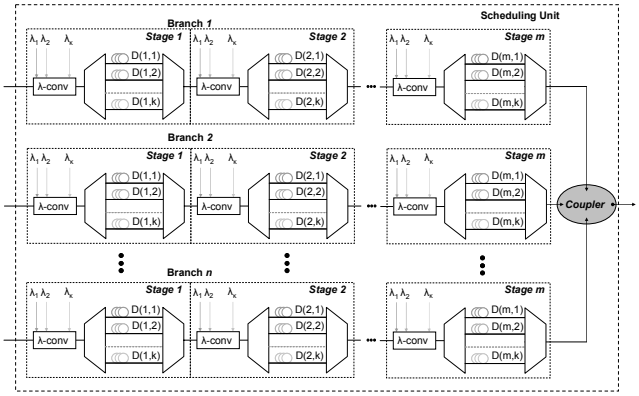

Fig. 3. Structure of the scheduling unit. It consists of *n* parallel delay branches, each comprising *m* serially connected programmable delay stages.

The algorithm is demonstrated in Fig. 2, for arbitrary length packet payloads arriving at a scheduling unit. As it is made clear from Fig. 2, the algorithm concentrates all the DUs that originate from the same incoming link *i*, by sequencing them according to their rank. Moreover, all the DUs of the incoming link *i* are delayed by the total number of DUs of all previous links.

The conentnion resolution algorithm is capable of performing without DU losses inside the scheduler provided that the *(n,T)* smoothness property holds, while in the same time successive DUs that comprise a packet/burst are not separated while traversing the scheduler, since they have successive ranks, thus packet/burst integrity is maintained in the node. Moreover, the order of packets and bursts that arrive from the same incoming link and exit through the same outgoing link is also maintained in the scheduler, and this feature is particularly important since packet or burst re-ordering is not required at the egress network routers.

B. Scheduler Structure

The scheduling algorithm requires that packet payloads can be delayed between zero and *2T-1* time slots. The latter case corresponds to placing a packet arriving from incoming link *n* at the last time slots of the outgoing frame. Time delay is inserted in the scheduling unit, the structure of which is illustrated in Fig. 3. As we show in the figure the scheduling units are formed from *n* parallel branches. Each branch contains *m* in total serially connected programmable delay stages, which in turn consist of *k* parallel delay units. Every delay unit is accessed by means of a wavelength converter and a wavelength selective coupler, similar to the space switch. The signals that drive the wavelength converters are generated in the control plane according to Eq. (2). Time delay is introduced in each DU by propagation through standard single mode fiber of length equal to (in time slots)

$$D(p,q) = (q-1)\cdot k^{\,p-1}, \quad p = 1,..., m, \quad q = 1,..., k \tag{3}$$

where *p* is the delay stage number and *q* is the delay unit number. The delay that is introduced lies anywhere between zero and k^m*-1*, thus the required number of stages is given by

$$m = \lceil \log_k (2T) \rceil \quad (4)$$

It follows from Eq. (4) that, for a given frame size T, the scheduling unit architecture allows for a trade-off between the number of wavelengths k and the number of delay stages m. This is particularly important when minimizing the cost of the node, since the number of the wavelength converters that operate in the scheduling unit is determined by m, while both k and m determine the required total fiber length of the delay units. The total fiber length required in a scheduling unit branch is given by

$$L = n^2 \cdot \sum_{p,q} D(p,q) = n^2 \cdot \frac{k \cdot \left(k^m - 1\right)}{2} \quad (5)$$

Due to the ceiling function in Eq. (4), the number of the required delay stages m may be the same for several numbers of wavelengths k. As a result, for a given number of stages, the minimum number of wavelengths, and therefore the minimum fiber length, should be chosen.

C. Scalability of the Node Core

The space-wavelength parallelism of the node core architecture, Spanke-like Switch and Scheduler, facilitates the expansion of the node size with respect to the number of its incoming and outgoing links. An extra incoming link is readily installed by adding a header extraction subsystem, a wavelength converter and a wavelength selective coupler. Moreover, an additional branch is required per scheduling unit and this can be achieved by cascading m delay stages, as shown in Fig. 3. An extra outgoing link, on the other hand, requires the deployment of a scheduling unit, a wavelength converter and a header re-insertion subsystem. In addition, the number of the output ports at the wavelength selective couplers of the Spanke-like switch must be increased. The replacement of the couplers can be avoided if couplers with redundant output ports have been pre-installed, as in Fig. 1. Therefore, the node core is easily scalable, and its extension requires only the deployment of additional components.

III. Wavelength Converter

The wavelength converters are key subsystems in the node architecture discussed earlier, since they perform incoming to outgoing link mapping and intra-node contention resolution. In this Section we discuss the structure and implications of operating all-optical wavelength converters that rely on optical gating. Optical gating takes place in a semiconductor optical amplifier (SOA) - based asymmetric Mach - Zehnder Interferometer (MZI) [11], which is illustrated in Fig. 5. Delayed replicas of the signal at the input wavelength are used to switch on and off the MZI switch, thus forming a switching window during which a continuous wave (CW) signal at the desired output wavelength is allowed to exit the MZI.

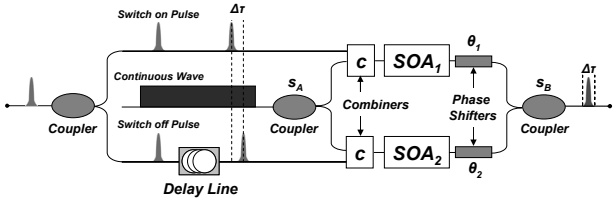

Fig. 5. Wavelength conversion in an asymmetric SOA based MZI. The input wavelength (signal) controls the switch on and off of the gate. The output wavelength (continuous wave) exits the gate during the switch on period.

Switching on of the MZI is achieved using the control pulses to induce gain and phase change on the CW signal in the upper arm SOA, whereas switching off is achieved in a similar fashion by inducing gain and phase change in the lower arm SOA.

The MZI switch on and off is governed by the following equations [11]:

$$P_{out}(t) = P_{in}^{CW} \cdot c^2 s_A s_b x_o^2 \cdot G_2(t) \cdot T\left(\frac{x(t)}{x_o}\right) \quad (6)$$

$$T\left[\frac{x(t)}{x_o}\right] = 1 - 2\frac{x(t)}{x_o}\cos\left\{a \ln\left[\frac{x(t)}{x_o}\right]\right\} + \left[\frac{x(t)}{x_o}\right]^2 \quad (7)$$

$$x(t) = \sqrt{\frac{G_1(t)}{G_2(t)}},\ x_o = \sqrt{\frac{G_1^{ss}}{G_2^{ss}}} \quad (8)$$

The parameters in Eq. (6) - (8) are: c is the cross wavelength coupling ratio in the combiners, s_A and s_B are the MZI input and output coupling ratios and a is the linewidth enhancement factor that correlates gain and phase changes according to $\phi(t) = -\frac{\alpha}{2}\ln[G(t)]$ [12]. The upper and lower branch SOA gains are denoted as $G_1(t)$ and $G_2(t)$ respectively, and G_1^{ss} and G_2^{ss} are the SOA small signal gains when no control pulses access them.

Eq. (7) describes the switching window of the MZI, and it follows that when the MZI operates without control pulses, both SOAs operate in the small signal gain regime and there is no output. On the contrary, when control pulses access the SOAs, they saturate the SOA gains, and output pulses are formed. However, due to the limited rate at which the SOAs recover their gain after saturation by the control pulses, data-pattern free operation of the wavelength converter is feasible only when the SOA gains are heavily saturated ($G_2(t) = 1$). As a result, the output power is of the same order of magnitude as the CW signal and there is an effective power loss when converting wavelengths, since the CW signal power is set to at least an order of magnitude less than that of the control signal power, so as to leave the switch dynamics unaffected. The power loss introduces a major difficulty when cascading the wavelength converters, and consequently realizing the optical node, since it affects the error-free

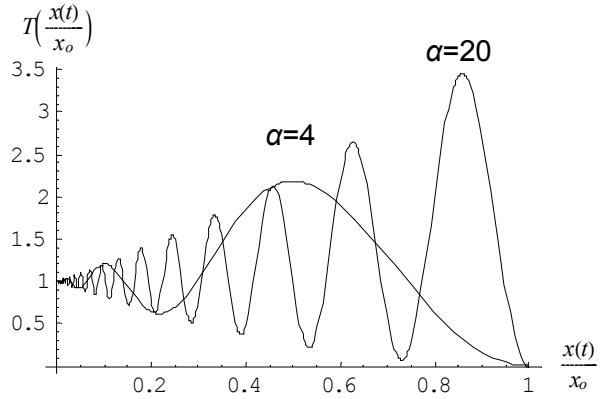

Fig. 6. Dependence of the SOA gain saturation on the linewidth enhancement a-factor. Wavelength conversion with light gain saturation is feasible for high *a*-factors.

operation of the wavelength converter at the immediately following stage, in which the converted signal of the current stage is required to control the switch on and off of the MZI.

Therefore, a major challenge when cascading MZI based wavelength converters is alleviating the power loss they exhibit. A possible solution is the deployment of inline amplifiers between wavelength converters, but this would impose an increase of the node cost. Alternatively, we propose operating the wavelength converters as both switching and gain elements. This is achieved, according to Eq. (6), when the SOA gains are slightly saturated ($G_2(t) \approx G_2^{ss}$). Data-pattern free operation is possible in the slight saturation regime since the SOAs fully recover their small signal gains within a bit period. However, slight saturation is possible only under the condition that the *a*-factor assumes large values. The effect of the *a*-factor on the SOA gain saturation is shown in Fig. 6, where the switching window is plotted against the normalized gain ratio $x(t)/x_o$ for a=4 and 20. Fig. 6 shows that higher *a* factor values force the maximum of the switching window to take place for gain ratios closer to x_o, therefore slighter gain saturation is required in the SOAs. Moreover, the switching window maximum itself reaches an upper bound that equals 4. As a result, the total gain that can be obtained in the slight gain saturation regime increases with the *a*-factor, and this enables the MZI to operate as both a switch and a gain element, contrary to the heavy saturation regime of operation. In addition, since the degree of SOA gain saturation depends on the power of the control pulse, slight saturation requires lower power levels of the control pulses. Therefore, even modest per MZI gains are expected to allow for the cascadability of the wavelength converters. We are currently investigating the impact of the *a*-factor on the MZI gain in order to fully implement cascades of wavelength converters.

IV. Conclusion

We have presented a lossless almost all-optical node architecture suitable for optical packet and burst switching. The node can be manufactured from commercially available equipment and its architecture is readily expandable, by simple deployment of additional components. The node core comprises a Spanke-like Switch, for connecting the incoming to the outgoing node links, and a Scheduler, for avoiding intra-node contention. Link mapping and contention resolution is based on the cascaded operation of wavelength converters, and we proposed a method for implementing a wavelength converter cascade without the deployment of additional in-line amplifiers.

References

[1] K. Fukuchi *et al.*, "10.92 Tbps (273x40 Gbps) triple-band/ultra-dense WDM optical-repeatered transmission experiment," *in Proc. Optical Fiber Communication Conference 2001*, pp. PD24.

[2] M. Nakazawa *et al.*, "1.28 Tbit/s-70km OTDM transmission using third- and fourth-order simultaneous dispersion compensation with a phase modulator," *Electron. Lett.*, Vol. 36, No. 24, pp. 2027-2029, Nov. 23 2000.

[3] Y. Muramoto et al., "Uni-travelling-carrier photodiode module with bandwidth of 80 GHz," *Electron. Lett.*, Vol. 39, No. 25, pp. 1851-1852, Dec. 11 2003.

[4] Y. Akage et al., "Wide bandwidth of over 50 GHz travelling-wave electrode electroabsorption modulator integrated DFB lasers," *Electron. Lett.*, Vol. 37, No. 5, pp. 299-300, Mar. 1 2001.

[5] P. Gambini *et al.*, "Transparent optical packet switching: network architecture and demonstrators in the KEOPS project," *IEEE J. Select. Areas Commun.*, Vol. 16, No. 7, pp. 1245-1259, Sept 1998.

[6] D.J. Blumenthal *et al.*" Optical signal processing for optical packet switching networks," *IEEE Communications Magazine*, Vol. 41, No. 2, pp. S23 - S29, Feb. 2003.

[7] M. Renaud *et al.*, "Network and system concepts for optical packet switching," *IEEE Communications Magazine*, Vol. 35, No. 4, pp. 96-102, April 1997.

[8] C. Qiao *et al.*, "The potentials of optical burst switching (OBS)," *in Proc. Optical Fiber Communication Conference 2003,* Vol. 1, pp. 219-220.

[9] C. Bintjas *et al.*, "All-optical packet address and payload separation," *IEEE Photon. Technol. Lett.*, Vol. 14, No. 12, pp. 1728-1730, Dec. 2002.

[10] E. Varvarigos, "The 'packing' and 'scheduling packet' switch architectures for almost all-optical lossless networks," *J. Lightwave Technol.*, Vol. 16, No. 10, pp. 1757-1767, October 1998.

[11] J. Leuthold *et al.*, "All-optical space switches with gain and principally ideal extinction ratios," *IEEE J. Quantum Electron.*, Vol. 34, No. 4, pp. 622-633, April 1998.

[12] G. P. Agrawal *et al.*, "Self-phase modulation and spectral broadening of optical pulses in semiconductor laser amplifiers," *IEEE J. Quantum Electron.*, Vol. 25, No. 11, pp. 2297-2306, Nov. 1989.